Spatio-temporal rainfall trends in southwest Western Australia

Ken Liang
University College London, ken@stats.ucl.ac.uk

Richard Chandler
University College London

Giampiero Marra
University College London

Abstract: Over the past several decades, there have been significant reductions in rainfall across southwest Western Australia. In the present work, the spatial and temporal structure of these reductions are investigated using generalized additive models. This involves smoothing over both space and time, to allow spatio-temporal interactions, as well as allowing for spatial correlation to ensure that standard errors are constructed appropriately for inference. The proposed method is computationally convenient as models are fitted as though different spatial locations are independent, and inference is subsequently adjusted for inter-site dependence. The results quantify precisely the spatially-varying nature of the decreasing rainfall trends.

Keywords: Spatio-temporal modelling, generalized additive models, tensor product smooth.

1 Introduction

Several decades of below average rainfall combined with a noticeable shift toward drier winter conditions, has focused attention on water resource availability and agricultural management in southwest Western Australia (SWWA) (Bates et al., 2008). The aim of this analysis is to characterize spatio-temporal trends in rainfall intensity and occurrence across SWWA. This is achieved by fitting generalized additive models (GAM) to data from selected locations in the study area. A key issue here is to take due account of potential spatial and temporal correlations in the data. Our approach to this is to treat the data as independent during fitting and subsequently to adjust standard errors for the dependence. This provides a computationally convenient means of addressing the problem.

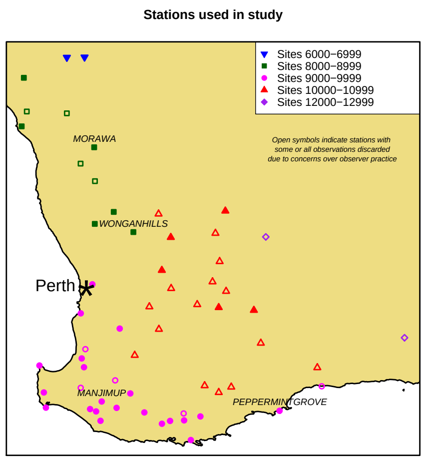

Figure 1: Map of SWWA indicating the stations used in the analysis

2 Materials and Methods

Daily rainfall readings in millimetres, from 60 weather stations (see Figure 1) for the period 1940-2010 have been used, although not all sites were operational throughout this period and some had missing observations. We consider the data for the winter months only, from May to July, which accounts for most of the region's annual rainfall. The daily rainfall data is aggregated in two ways, the proportion of wet days and total rainfall on wet days, for each site and year.

Prior to fitting the model, the daily rainfall data were subject to data quality checks. Issues such as inconsistencies in the data related to observer practice across the different sites, differences in the resolution of the observation recordings, and thresholding the data prior to analysis to ensure consistency have all been considered. These problems are typical in the rainfall modelling literature, the interested reader is referred to the results discussed in Yang et al. (2006) and Chandler et al. (2011). We shall not go into these details here.

Rainfall, particularly at the daily time scale, typically displays some form of temporal dependence, however at annual timescales they are relatively independent. Here, the winter rainfall is aggregated at an annual level, it seems reasonable to proceed in the first instance as though observations are independent between years. Since interest lies in characterizing temporal trends which may have a complex structure and may be spatially-varying, we adopt a nonparametric approach and represent the spatio-temporal trend surface as a smooth three-dimensional function of space and time:

$$E[y_{it}] = f(longitude_i, latitude_i, time_t),$$

where y_{it} is the aggregated annual winter rainfall at location $i = 1, \ldots, 60$ and year $t = 1, \ldots, 71$ and is assumed to be normally distributed. To account for different number of observations per year at each station, when fitting our model each observation is weighted by the number of contributing daily values. Our method uses the spline framework for nonparametric function estimation. To model smooths of several variables, when the variables are on different scales (the units of time (years) and space (km) are different), tensor product scale invariant smooths are required. Separate smoothing penalties are calculated for the three covariates so that the degree of smoothness is not necessarily the same for each covariate. All statistical analyses were done using the mgcv package (Wood, 2006) in the R software (R Development Core Team, 2010).

A key issue with this study is the spatially correlated nature of the data. Supposing that some assumption of spatial stationarity holds, the residuals from the fitted model were used to estimate the spatial correlation parameters obtained from the chosen variogram or correlation model, for the construction of the spatial correlation matrix. Specification of a covariance structure based on the spatial correlation of residuals ensures that the results are adjusted for spatial correlation. Because the focus of our approach is on estimating the mean function, not the correlation function, a very precise estimate of the latter is not required, and simple variogram models like the exponential will often suffice. The fitted function can be written as a linear smoother, $\hat{\mathbf{y}} = \mathbf{S}\mathbf{y}$ where $\mathbf{S}$ is the smoothing matrix. For spatially correlated observations, the true variance matrix is not diagonal. The model based variance matrix, $V(y)$, is replaced by the robust variance matrix, $\mathrm{Var}(y) = \mathbf{A}^{\mathbf{1/2}}\mathbf{R}\mathbf{A}^{\mathbf{1/2}}$ where $\mathbf{A}$ is a diagonal matrix, with the variance function $V(\mu)$, along diagonal elements and $\mathbf{R}$ is the spatial correlation matrix. Once the variance-covariance matrix is calculated, standard errors are then constructed in the usual way, as the square-roots of the elements on the main diagonal.

3 Results

After addressing spatial correlations, new 95% uncertainty bands were obtained which are now slightly wider than the unadjusted ones (see Figure 2). Interestingly, the width of the bands for the selected four sites differ significantly. This could be due the sparseness of data points at particular locations within SWWA which makes it difficult to characterize precipitation trends reliably. These declines were most pronounced in north and east of the study region, less so along the south coast.

4 Concluding remarks

To sum up, the used GAM framework enables us to appropriately incorporate all relevant covariates of space and time. In addition, accounting for the spatial dependence structure by assuming no correlation when fitting models and then adjusting

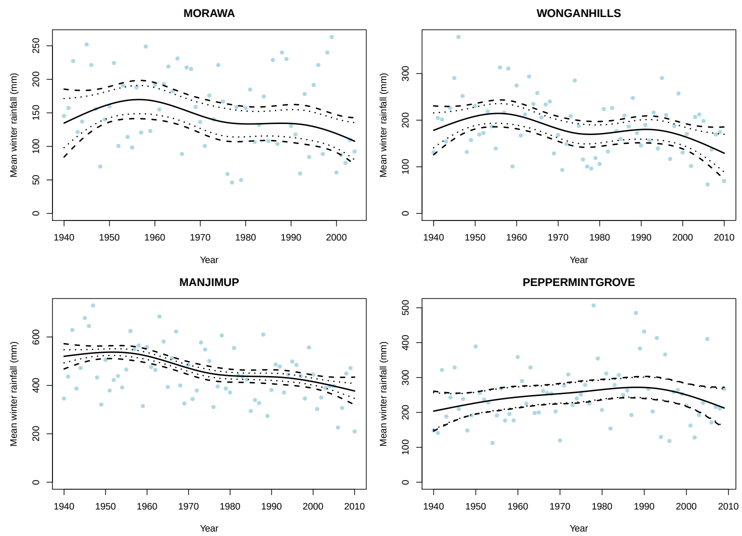

Figure 2: Fitted smooth curves (solid line) with unadjusted (dotted line) and adjusted (dashed line) 95% uncertainty bands for four selected sites in the SWWA region

the standard errors of estimates enable valid inferences that is robust. The results quantify precisely the spatially-varying nature of the decreasing rainfall trends.

References

Bates B., Hope P., Ryan B., Smith I. & Charles S. (2008) Key findings from the Indian Ocean Climate Initiative and their impact on policy development in Australia, *Climatic Change*, 89, 339-354.

Chandler R. E., Bates B. C. & Charles S. P. (2011) Rainfall trends in southwest Western Australia, in: *Statistical Methods for Trend Detection in the Environmental Sciences*, In Chandler R. E. & E. M. Scott (Eds.), Wiley, Chichester, 307-332.

R Development Core Team (2010) *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing, ISBN 3-900051-07-0.

Wood, S. (2006) *Generalized Additive Models. An Introduction With R*, Boca Raton: Chapman & Hall/CRC.

Yang C., Chandler R. E., Isham V. S. & Wheater H. S. (2006) Quality control for daily observational rainfall series in the UK, *Water and Environment Journal*, 20(3), 185-193.